Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism

Matteo T. Degiacomi[1#†], Ioan Iacovache[2#‡], Lucile Pernot[2], Mohamed Chami[3], Misha Kudryashev[3], Henning Stahlberg[3], F. Gisou van der Goot[2*] and Matteo Dal Peraro[1*]

[1] Institute of Bioengineering and [2] Global Health Institute
School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne - EPFL
CH-1015 Lausanne, Switzerland, [3] Center for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum, University Basel, Mattenstrasse 26, CH-4058 Basel, Switzerland

[#] These authors contributed equally to this work

* Correspondence to be addressed to G.F.V.D.G. or M.D.P. (co-senior authors)

Institute of Bioengineering / Institute of Global Health
Ecole Polytechnique Fédérale de Lausanne
Station 19, 1015 Lausanne, Switzerland
E-mails: matteo.dalperaro@epfl.ch, gisou.vandergoot@epfl.ch

[†] Current address: Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford, OX1 3QZ
[‡] Current address: Institute of Anatomy, University of Bern, Baltzerstrasse 2, 3012 Bern

Abstract

Aerolysin is the founding member of a super-family of ß-pore forming toxins for which the pore structure is unknown. We have combined X-ray crystallography, cryo-electron microscopy (EM), molecular dynamics and computational modeling to determine the structures of aerolysin mutants in their monomeric and heptameric forms, trapped at various stages of the pore formation process. A dynamic modeling approach based on swarm intelligence was applied whereby the intrinsic flexibility of aerolysin extracted from new X-ray structures was utilized to fully exploit the cryo-EM spatial restraints. Using this integrated strategy, we obtained a radically new arrangement of the prepore conformation and a near-atomistic structure of the aerolysin pore, which is fully consistent with all biochemical data available so far. Upon transition from the prepore to pore, the aerolysin heptamer shows a unique concerted swirling movement, accompanied by a vertical collapse of the complex, ultimately leading to the insertion of a transmembrane ß-barrel.

INTRODUCTION

Pore forming proteins (PFPs) are produced by a variety of organisms and are often involved in defense or attack mechanisms[1]. One striking feature is that they are produced as soluble proteins which subsequently oligomerize and convert to a transmembrane pore in the target membrane[2]. The most extensively characterized PFPs are the bacterial pore forming toxins (PFTs), which, depending on the secondary structure elements that cross the bilayer, have been classified as α- or ß-PFTs[2]. Aerolysin, produced by *Aeromonas sp.,* is the founding member of a large super-family that spans all kingdoms of life[3]. While it was the first ß-PFT for which the X-ray structure of the soluble form was solved[4], the structure of the pore has remained elusive.

Aerolysin is produced as an inactive precursor, proaerolysin, which contains a C-terminal peptide (CTP) required for folding into its soluble form[5]. Proteolysis in the loop that connects the CTP to the main body allows aerolysin to oligomerize in a heptameric ring-like complex that inserts into the target membrane to form the pore. Proaerolysin is an L-shaped molecule (Fig. 1a), where domain 1, involved in binding N-linked oligosaccharides,[6] is connected by a flexible arm to an elongated ß-rich lobe divided into three domains: Domain 2 involved in binding the glycan core of the Glycosyl Phosphatidyl Inositol (GPI)-anchored aerolysin receptors, domain 3 involved in oligomerization, and domain 4 that contains the CTP.

Early electron microscopy (EM) studies of 2D crystals of aerolysin pores showed heptameric ring-like structures[7]. When the proaerolysin structure was solved[4], monomers were rigidly docked into a mushroom-shaped 3D EM map with domains 1 and 2 alternating in the mushroom cap, and domains 3 and 4 forming the stem, thus generating the first model of the aerolysin pore. Based on the exposure of a hydrophobic pocket upon removal of the CTP, domain 4 was proposed to be transmembrane, positioning the mushroom stem in the bilayer. This initial model was subsequently refined with the study of an aerolysin point

mutant (i.e. Y221G), which forms soluble heptamers suitable for single particle EM analysis[8].

This model was first challenged when GPI anchored proteins were shown to be the aerolysin receptors[9,10]. Domain 1 binds N-linked sugars on the receptor (Fig. 1a: residues K66, I47, M57, Y61 and W45)[6,11], while domain 2 binds the glycan core of the GPI-anchor (Fig. 1a: residues H332, W324 and Y162)[11], i.e. to the lipid head group. In the early aerolysin pore model, receptor binding sites in domain 2 are positioned >10 nm above the membrane surface. Such a distance would argue for receptor release upon heptamerization, which we know does not occur since Y221G heptamers remain membrane bound, without membrane insertion[12]. Thus domain 2 cannot be positioned 10 nm above the membrane.

The second challenge came from the identification of residues crossing the membrane[13]. We showed that an amphipathic loop, initially sandwiched between domains 3 and 4, undergoes sliding, realignment and association with the corresponding structure in neighboring monomers, generating an amphipathic 14-stranded ß-barrel that spontaneously inserts into the bilayer[13], as also found for the related toxin, *Clostridium septicum* α-toxin[14]. This loop (hereafter called pre-stem loop, from residues 246 to 258, inset in Fig. 1a). forming the transmembrane portion is again incompatible with the initial model, where domain 4 is transmembrane[4,8].

We here propose a novel paradigm for aerolysin assembly and pore formation. Using new X-ray structures, cryo-EM maps of aerolysin mutants trapped at intermediate states along the pore formation process, molecular dynamics and swarm intelligence optimization searching, we build and validate near-atomistic models of aerolysin at different stages of pore formation. Our models reveal a unique membrane insertion mechanism, involving a remarkable conformational switch upon transition from prepore to pore, wherein the aerolysin heptamer shows a concerted swirling movement, accompanied by a vertical collapse of the complex, upon membrane insertion. The drastic conformational

changes involved in aerolysin pore formation bare no similarity to those described for other PFTs such as *Staphylococcus* a-hemolysin or the cholesterol dependent toxin perfringolysin O[2].

RESULTS

Whereas the pro-aerolysin structure is available, crystallization attempts of the pore have failed, possibly due to high protein flexibility. We therefore opted for a detour strategy combining X-ray crystallography, cryo-EM, single particle analysis, molecular simulation and modeling. Similar integrative approaches have been successful to understand multi-protein complex formation[15-17].

Wild-type aerolysin heptamers are not amenable to single particle analysis because they aggregate into heterogeneous clusters. Instead, we analyzed two previously characterized mutants that retain the ability to form heptamers, but are blocked at intermediate stages of the pore formation process. The first, Y221G (Fig. 1), is blocked at the prepore stage and thus forms soluble heptamers[8]. The second, K246C–E258C (inset in Fig. 1a), contains an engineered disulfide bridge that does not hinder heptamerization, but prevents the complete realignment of the pre-stem loop into a ß-hairpin consequently inhibiting pore formation. Importantly, K246C–E258C heptamers are partly hydrophobic, as opposed to Y221G heptamers, indicating that they are blocked beyond the prepore stage[8].

Aerolysin is intrinsically flexible upon activation

The ability of proteins to assemble into multimeric complexes is frequently related to their capacity to explore multiple states within their conformational space. Consistently, large conformational changes were reported for the PFT *E. coli* ClyA[18] and perfringolysin[19]. To obtain clues regarding the possible molecular

rearrangements associated with heptamerization, we characterized the intrinsic flexibility of aerolysin based on the X-ray structure of WT[4] and Y221G, here determined at 2.2 Å resolution (showing a backbone root mean square deviation of 1.3 Å from the WT, Supplementary Results, Supplementary Fig. 1a and Supplementary Table 1).

Molecular dynamics (MD) simulations on a 10^2 ns timescale were stable for both WT and Y221G pro-aerolysins and sampled a fairly limited conformational space (Supplementary Fig. 1b). Upon CTP removal, corresponding to the activated state, both WT and Y221G species explored a broader conformational space, consistent with our observations that the CTP maintains folding of domain 4[5]. In particular, flexibility of domain 4 with respect to the rest of the protein was increased. Flexibility was far more pronounced for WT than Y221G, revealing different dynamic signatures in the activated state (Fig. 1b and Supplementary Fig. 1).

A new membrane orientation for the Y221G prepore

We previously analyzed Y221G soluble heptamers by single particle cryo-EM on negatively stained samples[8]. Here we performed cryo-EM on Y221G heptamers without stain to obtain a more accurate model. As previously, Y221G associates into a symmetric head-to-head dimer of heptamers (Fig. 2a). Each heptamer shows a mushroom-like shape, with a disc-like cap region and a stumpy stalk. Single-particle averaging led to a 16.6 Å resolution EM map of the Y221G prepore (Fig. 2a,b and Supplementary Fig. 2a,c).

To model the Y221G prepore, we developed an optimization framework based on swarm intelligence heuristic[20] that integrates the intrinsic protein flexibility explored by MD with the spatial restraints imposed by the cryo-EM map. Using such approach an ensemble of physical-based structural conformations can efficiently probe assembly states consistent with a set of spatial restraints.

The power of current MD sampling is crucial to explore at near-physiological conditions the broader conformational space of single subunits upon assembly. For Y221G aerolysin, MD sampling not only revealed greater flexibility when compared to the X-ray structure, but *in silico* mimicking of the activation process through removal the CTP highlighted previously unrecognized fluctuation modes of domain 4 (Fig. 1b and Supplementary Fig. 1b). Following this assembly protocol (see Methods), the global dimensions of the EM map were used to assemble the prepore considering the ensemble of Y221G structures produced by MD. The best heptameric model had a cross-correlation coefficient (CCC) with the cryo-EM map of 0.79; after refinement using a MD flexible fitting procedure[21], the final structure optimally fitted the experimental data with a CCC=0.87 (Fig. 2b). It is noteworthy that, when the static X-ray structure of Y221G aerolysin was used, none of the explored solutions satisfied the spatial restraints.[20] Therefore, intrinsic protein flexibility is essential when performing docking into EM maps of multi-subunit complexes[15,22]. In the specific case of aerolysin, flexibility was only revealed upon *in silico* CTP removal, a step known to be necessary for heptameric assembly[5].

Inspection of the Y221G prepore model reveals that the GPI-anchor binding sites are located at the edges of the mushroom cap, indicating that this surface must be in close proximity to the lipid head groups (Fig. 2c and Supplementary Fig. 3b). Placing the mushroom cap towards the membrane surface positions the stalk upwards into the extracellular milieu. Therefore, the orientation we propose corresponds to a 180° rotation compared to the initial aerolysin pore model[4,8].

Domain 2 binds to the glycan core of GPI-anchors, which share the same backbone structure, ethanolamine-HPO_4-6Manα1-2Manα1-6Manα1-4GlcNH2α1-6-myo-inositol-1HPO_4, linked to the lipid moiety. Aerolysin can bind mannose residues of this backbone in the absence of any modification[23]. To gain structural information on binding of aerolysin to GPI-anchors, we soaked Y221G pro-

aerolysin crystals with sugars and determined the X-ray structure of Y221G pro-aerolysin at 2.5 Å resolution with a mannose molecule trapped in domain 2 (Supplementary Fig. 3a and Supplementary Table 1). We next generated a model of the smallest known GPI-anchored protein, CD52[24], which is composed of only 12 amino acids including an N-glycosylation site, and which is an aerolysin receptor[23] (see Methods). Superimposition of the GPI-anchored CD52 embedded in a membrane bilayer on the relative position of the mannose moiety trapped in the X-ray structure confirmed the orientation of the Y221G prepore with respect to the membrane (Fig. 2c and Supplementary Fig. 3b). Altogether, these observations indicate that domain 2 in the prepore is facing the membrane and that domain 3 acts as a hinge that positions domain 4 upwards.

To validate the atomistic accuracy of the prepore structure, residues located at the monomer-monomer interfaces were experimentally mutated to alanine and the ability of the mutants to form heptamers *in vitro* was compared to that of WT. Mutants K198A, D216A, R282A, corresponding to domain 4 interfaces, K369A, E367A at domain 2-domain 1 interfaces and K351A at the domain 2-domain 3 interface had slower kinetics of oligomerization, heptamerization being almost blocked for certain mutants, such as D216A (Fig. 2d and Supplementary Fig. 4).

The Y221G mutation prevents pre-stem loop release

The major difference between WT and Y221G aerolysin is that Y221G, although able to form heptamers, fails to form the transmembrane ß-barrel. As mentioned, the ß-barrel is formed by the pre-stem loop (i.e. segment K246-E258, Fig. 1a). Our model of the Y221G prepore shows that the pre-stem loop lies in the vicinity of an area of low density (apparent hole) in the EM map (Fig. 2b, inset). This location suggests that in the WT protein, the pre-stem loop might slide through the hole into the center of the heptameric ring, and thus be optimally positioned to assemble, with neighboring loops, into the transmembrane ß-barrel.

We investigated whether the comparison of WT and Y221G structures would provide a clue as to why extraction of the pre-stem loop from its resting position would be prevented by the Y221G mutation. MD simulation on the WT shows that, upon CTP removal, domain 4 starts to unfold, the most affected strands spanning residues L219-K229 and N269-V281 (Fig. 3a). Interestingly, unfolding is initiated in the neighborhood of residue Y221 and propagates towards the pre-stem loop (Fig. 3b). In marked contrast, CTP removal has only a moderate effect on Y221G secondary structure. Indeed, after similar simulation times, the original secondary structure is largely preserved and completely conserved in the strand region approaching the pre-stem loop. Consistently, in the Y221G pro-aerolysin X-ray structure (Supplementary Fig. 1a) the presence of a glycine at position 221 leads to the flipping of L277 into the position of Y221 in WT (inset in Fig. 3a). Both analyses indicate that L277 maintains this position at the G221 pocket, and neighboring residues preserve their initial ß-sheet secondary structure with and without the CTP in place. Therefore, in the absence of a tyrosine at position 221, L277 critically maintains order and stability in domain 4. These observations suggest that in WT, due to its peculiar position in the hydrophobic pocket underneath the CTP, movement of Y221 upon CTP removal acts as a trigger for the conformational change in domain 4, which subsequently propagates to the pre-stem loop to slide away from its resting position. This is also consistent with our previous finding[5] that release of the CTP leads to unfolding of the two ß-strands that feed into the pre-stem loop.

Pre-stem loop extraction increases aerolysin flexibility

The next key step towards pore formation is the extraction of the pre-stem loop. We observed that steering out the pre-stem loop in MD leads to a remarkable rearrangement of domain 4, which heavily reorients with respect to domain 3 (Supplementary Fig. 5). By considering the aerolysin domains as rigid bodies, domain 4 of pro-aerolysin fluctuates around an equilibrium state that closely

resembles the arrangement of the crystal structure (Fig. 1a,b and Supplementary Fig. 1). However, upon pre-stem loop extraction and CTP removal, domain 4 twists sidewise and completely flattens with respect to domain 3 (Supplementary Fig. 5), drastically affecting the tertiary conformation of aerolysin, and hinting to major effects on quaternary pore structure.

Snapshots of the pore *en route* to membrane insertion

To gain information on the conformational changes occurring beyond the prepore stage, we analyzed K246C–E258C heptamers by EM. This mutant forms heptamers that show hydrophobicity but fail to complete pore formation, due to the disulfide bond within the pre-stem loop that prevents proper ß-hairpin formation[13]. EM imaging revealed the formation of homogeneous complexes amenable to single particle analysis. Class average 3D reconstruction led to an 18.3 Å resolution EM map (Supplementary Fig. 2b,c). K246C–E258C particles were composed of dimers of heptamers (Fig. 4a), but, the two heptamers had clearly different conformations (Fig. 4a,b), and close inspection revealed that both were different from those of the Y221G prepore (Fig. 2b).

One K246C–E258C heptamer resembles Y221G heptamers (Figs. 4b,c and 2b), but shows a change in the vertical protrusion, which appears more cylindrical (Fig. 4b,c). Since this portion corresponds to domains 3 and 4 (Fig. 2b), we speculated that this K246C–E258C configuration corresponds to a *post-prepore* state, where the pre-stem loop had moved away from its resting position (Fig. 4c). Importantly, a density is observed in the internal cavity of the post-prepore heptamer (Fig. 4c, middle panel), which is absent in the prepore (Fig. 4c, top panel) and might correspond to the pre-stem loop.

The second heptamer in the K246C–E258C particles has a disk-like structure lacking the protrusion (stalk of the mushroom) observed in both prepores and post-prepores (Fig. 4c, bottom panel). A stem structure is apparent

between the two asymmetric heptamers on the raw images and class averages (Fig. 4a, red arrow), although the flexibility and low contrast of the stem structure result with its loss in the 3D reconstruction due to density averaging. This second conformation of the K246C–E258C heptamer was termed *quasi-pore*, since full completion of ß-barrel formation is prevented by the C246–C258 disulfide bridge. Interestingly, the hole observed in vicinity of the pre-stem loop in the Y221G prepore (Fig. 2b) and still present in the post-prepore (Fig. 4c and Supplementary Fig. 6) is absent in the quasi-pore K246C–E258C cryo-EM map (Fig. 4b,c bottom panel). It appears that sealing of the hole by the sliding of the stem loop leads to an increase in the contact surfaces of adjacent monomers, possibly contributing to the extraordinary stability of the aerolysin WT heptamer[25].

The K246C–E258C asymmetric dimers of heptamers are quite extraordinary structures where one heptamer (i.e. quasi-pore) attempts to form a pore in the central cavity of the adjacent heptamer (i.e. post-prepore). Once one heptamer has undergone partial pore formation, it prevents, probably due to steric hindrance within the central cavity, the facing heptamer from undergoing a similarly conformational change. The pre-stem loop in this latter conformation can however still move away from its resting position, altogether resulting in the formation of quasi-pore/post-prepore dimers (Supplementary Fig. 6).

As mentioned, WT heptamers aggregate and are not amenable to single particles analysis, probably because full formation of the ß-barrel forces the heptamers in the dimer to separate leading to the exposure of hydrophobic surfaces. Nonetheless, upon centrifugation of WT heptamer samples and EM analysis of the supernatant, a few heptameric dimers could be distinguished (Fig. 4d). As for K246C–E258C, these adopted different conformations, one heptamer showing a mushroom-like shape and the other a disk-like structure. Disk-like structures were also observed upon reconstruction of WT heptamers in nano-disks[26] (Supplementary Fig. 7a). Altogether, these experimental and previous computational (Supplementary Fig. 5) observations indicate that also for WT a

progressive flattening of the heptameric structure occurs during prepore-pore transition.

The EM density maps obtained for Y221G and K246C–E258C thus provide three snapshots of the aerolysin heptamer on its way to pore formation (Figs. 2 and 4), where the quasi-pore structure reaches a conformation close to that of the membrane-inserted WT pore.

Near-atomistic model structures of the aerolysin pore

Previously we established that residues W^{247}PLVG form the tip of the membrane-spanning ß-hairpins and that at least 7 residues of each side compose the ß-strands[13]. Using the same combination of cysteine-scanning mutagenesis and pore analysis in planar lipid bilayers, we here identify the boundaries of the ß-barrel. N226, T230 and G234 when mutated to cysteine show channel closure upon labeling with cysteine reactive biotin (Supplementary Fig. 8a) indicating that these residues are in the pore lumen. Taking residues K229 to N269 as the boundaries, we next built the ß-barrel using, as a guide, the spatial restraints extracted from previous[13] and current cys-scanning data, and the scaffold of the *S. aureus* hemolysin ß-barrel, whose conformation is likely shared with aerolysin[27] (Fig. 5a and Supplementary Fig. 8b).

The extra-membrane portion of the pore was modeled using K246C–E258C density map of the quasi-pore (Fig. 4b,c) and the flexibility observed in aerolysin upon CTP removal and stem displacement (Supplementary Fig. 5). As for the prepore state, we used swarm intelligence optimization to produce an ensemble of models fitting the global shape of the EM density, the best of which showed a CCC of 0.73, which was further refinement to 0.90 using MD flexible fitting[21]. We then linked S228 and G270 of the extra-membrane structure to K229 and N269 of the ß-barrel (Fig. 5a), obtaining thus the first near-atomistic view of the architecture of the membrane-inserted aerolysin pore (Fig. 5b).

The model reveals a large rearrangement of (i) the monomer unit when compared to the static X-ray structures of pro-aerolysin, and (ii) of the heptameric conformation when compared to pre-pore. In particular, in the pore, domain 4 of one monomer lies on parallel to the membrane (Fig. 6). This is in marked contrast to the prepore in which domain 4 is aligned almost perpendicularly with respect to the lipid bilayer (Fig. 2b). Domain 4 thus needs, according to its intrinsic flexibility (Fig. 1b and Supplementary Figs. 1b and 5) to twist sideways in a swirling fashion during prepore-pore transition (Supplementary Video 1). The comparison of the prepore and pore models (Figs. 2b, 5b and 6) shows that the transition from one conformation to the other is possible without clashes and topological bottlenecks. Importantly, this pore conformation is also structurally consistent with low-resolution images of the aerolysin pore entrapped in nano-disks (Supplementary Fig. 7) and observed by cryo-electron tomography (data not shown).

DISCUSSION

While the structure of pro-aerolysin was solved in 1994[4], that of the membrane-inserted oligomer remained elusive. Combining X-ray crystallography, cryo-EM reconstruction, mutagenesis, molecular simulation and modeling we generated structural models of aerolysin heptamers at various stages of the pore formation process. Remarkably, these models are consistent with all available structural and biochemical information regarding the aerolysin mode of action, and contribute to the understanding of some unexplained observations. Our work reveals a new paradigm of pore formation by members of the aerolysin superfamily that is drastically different from that of the two other classes of ß-PFTs[2].

Pro-aerolysin monomers approach the host membrane binding the GPI-anchor and N-linked sugars of its receptors. The restrained position of the GPI-

anchor on the lipid head group orients domain 2, and thereby the entire pro-aerolysin molecule with respect to the membrane. In particular, this conformation positions the CTP cleavage loop in domain 4 favorably for recognition by activating proteases in the extracellular space. Following cleavage, monomers assemble into heptamers, a step that requires CTP release. The present models show that in the heptamer, the hydrophobic pocket initially masked by the CTP is completely covered by a neighboring monomer (Figs. 2b and 5b), raising the possibility that CTPs are removed during heptamerization by competitive binding of incoming adjacent subunits.

Monomers first assemble into a transient prepore, which is trapped by the Y221G mutation. In the prepore, the GPI-binding sites remain favorably exposed (Fig. 2), consistent with the experimental observation that aerolysin remains receptor-bound following oligomerization[12]. The model also explains why residues G95 to L134 of *C. septicum* α-toxin were found to interact with the membrane[28]. The corresponding region in aerolysin, G137 to K167, is indeed located in domain 2 between the complex and the lipid bilayer, both in the prepore and the pore configurations (Figs. 2b and 5b). The membrane-facing surface of the cap does not show any extended hydrophobic patches, suggesting that it does not deeply interact with the bilayer in contrast to what is observed for α-hemolysin[27] or cytolysin[29]. The absence of hydrophobicity is consistent with the observations that aerolysin pores are somewhat distant from the host membrane surface[7] as can be seen in nano-disk entrapped pores (Supplementary Fig. 7a).

The prepore-pore transition involves a spectacular rearrangement of the extra-membranous region of the aerolysin heptamer, which converts from an inverted mushroom shape to a disk-like structure (Fig. 6, Supplementary Video 1). This massive change is in marked contrast to what is observed for *S. aureus* hemolysin, where the cap domain remains relatively unchanged when the transmembrane ß-barrel forms[2]. Also, despite a significant vertical collapse upon

membrane insertion, little change in tertiary structure has been reported upon pore formation by cholesterol-dependent cytolysins[2].

The pronounced preopore-pore rearrangement is consistent with previous studies showing significant differences in tertiary, but not secondary, structure between the Y221G and WT heptamers[8]. The proposed mechanistic transition is further supported by the observation that mutations of residues predicted to lie at the monomer-monomer interface in both the prepore and the pore (K198, D216, R282, K369, E367 and K351) do affect the capacity of aerolysin to oligomerize (Fig. 2d and Supplementary Fig. 4a), while mutations of residues that are found at protein-protein interfaces in the prepore, but are predicted to be solvent exposed in the pore (K185, D188, K290, E307 and K309), do not affect heptamerization significantly (Supplementary Fig. 4b).

The pre-stem loop reveals an extraordinary dynamics: initially positioned in a pocket, from which it is gradually extracted to slide through a cavity between two monomers, it ultimately extends and realigns into a ß-hairpin, which associates with the neighboring hairpins into a barrel (Fig. 6a and Supplementary Fig. 8b). Following extraction from its resting position, the pre-stem loop could be guided towards the membrane by a large Ser/Thr track (T223-T232 and S272-S276) that is found at the transition region between the barrel and the cap. Ser/Thr tracks have also been reported in parasporin[30] and lysenin[31], two members of the aerolysin super-family, suggesting that they are essential for pore formation.

The ß-barrel structure, based on cysteine-scanning mutagenesis, shows that the length matches that of the *S. aureus* α-hemolysin pore[27]. The alternating hydrophobic-hydrophilic pattern, the hallmark of transmembrane ß-barrels, is present over ~25 Å (Fig. 5a and Supplementary Fig. 8b), matching the hydrophobic region of an average eukaryotic bilayer. The alternating pattern terminates at a ring of outwards facing aromatic residues formed by W265 and Y233. This aromatic belt provides an ideal anchor at the external boundary of the

bilayer. Another striking feature of the aerolysin ß-barrel is the large number of charged residues pointing into the lumen (from position 198 to 290, for a total of 91 charges organized in rings, of which 7 positive unpaired). This is again in marked contrast to members of the *S. aureus* PFTs, which lack charged residues in the pore lumen. The two toxin families should therefore have significantly different ion selectivity, in particular with respect to Ca^{2+}, a key ion in terms of cellular responses to PFTs[32].

In conclusion, we propose new near-atomistic models of the prepore and transmembrane pore of aerolysin that shed light on the mechanism of pore formation, reconciling all available information on aerolysin mode of action. Our work reveals a swirling mechanism for membrane insertion which is likely to be shared by the other members of the aerolysin super-family[3], but is fundamentally different from the mechanism of membrane insertion of other PFTs. The integration of experimentally-derived restraints with the conformational ensemble characterized by molecular dynamics sampling was key in revealing the architecture of aerolysin pores; this approach[20] will likely be useful for the dynamic modeling of other large molecular machines.

Acknowledgements

We thank Fabrice David and Sylvia Ho for technical assistance, Graham Knot and Davide Demurtas for help with EM, Markus Grütter for testing of crystals, Clemens Schulze-Briese for assistance in collecting the data, Marc Schiltz for support with X-ray crystallography. EM computation was performed at the Vital-IT (Swiss Institute of Bioinformatics). This work was supported by the Swiss National Science Foundation with grants to M.D.P. (200021_122120 and 200020_138013) and to F.G.v.d.G., and by the Swiss Initiative for Systems Biology, SystemsX.ch to F.G.v.d.G. and H.S.

Author contributions

F.G.v.d.G. and M.D.P. designed and supervised the study. M.T.D. performed simulations and modeling. I.I. performed cryo-EM analysis, molecular biology, nanodics and biochemical experiments. L.P. performed protein production and X-ray crystallography. M.K, M.C. H.S. supported the EM. M.T.D., I.I., F.G.v.d.G. and M.D.P. wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available in the online version of the paper. Correspondence and requests for materials should be addressed to F.G.v.d.G. or M.D.P.

Figure legends

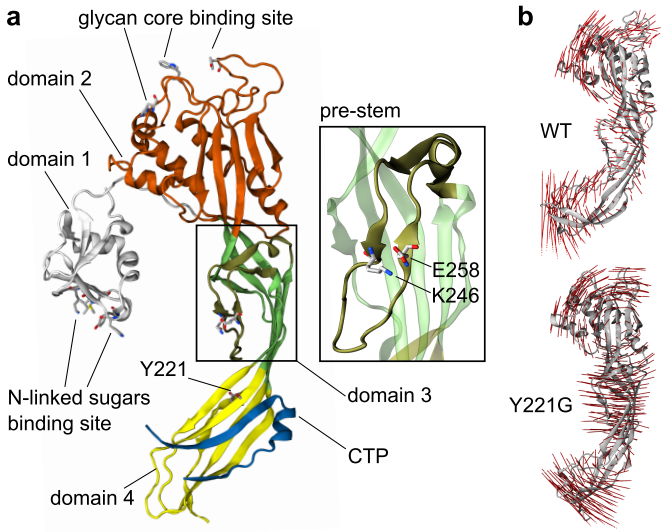


Figure 1. Aerolysin structure and flexibility. (a) Structure of monomeric pro-aerolysin (pdb 1PRE). Structural domains are color coded: domain 1 in gray, domain 2 in orange, domain 3 in green and domain 4 in yellow. The C-terminal peptide (CTP) is shown in blue and the pre-stem domain, enlarged in the inset, is shown in tan. Residues involved in binding to the receptor are highlighted on domain 1 (binding to N-glycans on the receptors) and 2 (binding to the glycan core of GPI-anchored receptors). Position of relevant mutations are highlighted in sticks: Y221, mutated to glycine produces a prepore state; cross-linking of pre-stem residues K246 and E258 (inset) produces a quasi-pore conformation. (b) Comparison of the flexibility of WT and Y221G aerolysin. The first eigenvector extracted by a principal component analysis of MD simulations is projected on Cα atoms of the MD averaged structures. Their relative length and direction quantify the entity of domain flexibility, when the CTP is removed from its position at domain 4 (see Supplementary Fig. 1b).

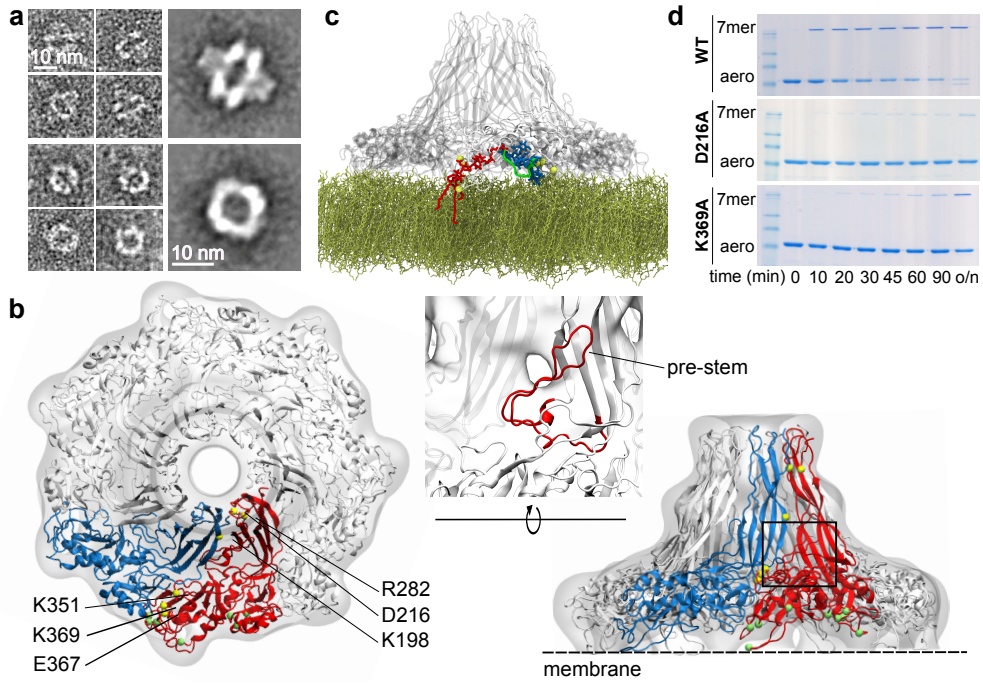


Figure 2. The prepore state trapped by the Y221G mutation. (a) Examples of side (top) and top (bottom) EM views Y221G particles as seen by negative staining (left) and their respective class averages (right) (see Fig. 2a,c). (b) Top and side views of the docking of seven Y221G monomers into the EM map at 16.6 Å resolution (Supplementary Fig. 2a,c). The position of key interface residues is shown in yellow space-filled representation. The positions of residues involved in receptor binding are shown in green space-filled representation. In the inset, the pre-stem loop (red) is adjacent to the low electron density (apparent hole) observed in the EM map. (c) Modeling of aerolysin prepore onto the GPI-anchored protein CD52. Domain 1 binds to N-linked sugars (in blue) located on CD52, while domain 2 binds to a mannose on the GPI-anchor (red) conserved core. The X-ray structure of pro-aerolysin complexed with a mannose-6-phosphate molecule is used to generate the complex (Supplementary Fig. 3a and Supplementary Table 1). The location of both binding sites forces the current orientation the prepore heptamer at the membrane (in yellow) (see Supplementary Fig. 3b). (d) *In vitro* kinetics of heptamerization of WT aerolysin and selected monomer–monomer interface mutants indicated in (b). Aerolysin

(WT and mutants) was activated at 4 °C to prevent oligomerization with insoluble trypsin. After removal of trypsin the sample was shifted at room temperature and heptamerization was monitored by SDS-PAGE. The complete data for all the interface residues are reported in Supplementary Fig. 4

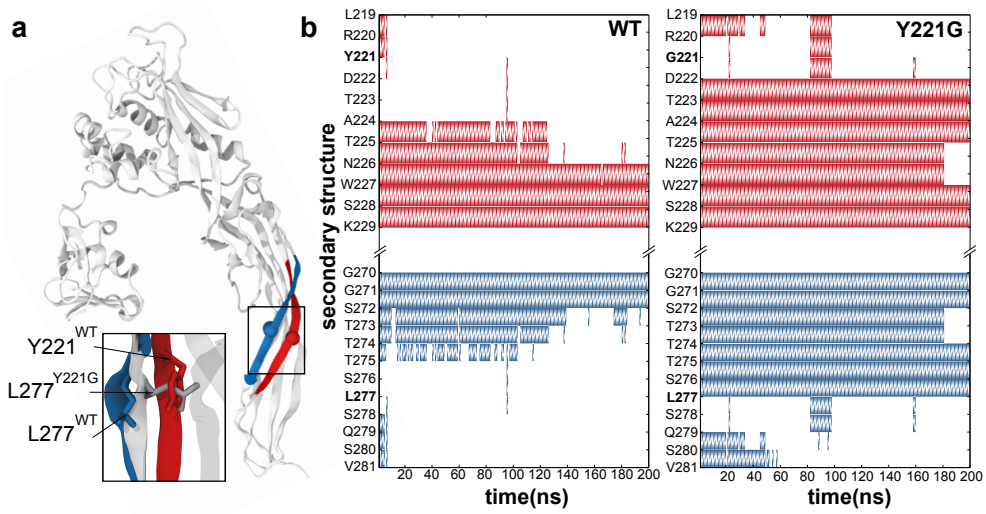


Figure 3. Pre-stem loop is locked in Y221G aerolysin. (a) X-ray structures of monomeric pro-aerolysin in WT and Y221G conformation are superimposed (see Supplementary Fig. 1a). Inset shows the different conformation adopted by residue L277 in either the WT (blue) or the Y221G (gray) X-ray structures. The two β-strands connecting domains 3 and 4 to the pre-stem loop are shown with strand L219 – K229 (in red) and N269 – V281 (in blue). Two spheres highlight the position of L277 (blue) and Y221 (red). (b) Evolution of the secondary structure during MD simulations of the WT (left) and Y221G (right) protein in the absence of the CTP. Red and blue colored regions represent conservation of β-strand structure indicated in (a). Upon removal of the CTP unfolding propagates throughout the two highlighted β-strands towards the pre-stem domain in WT, but stops at G221-L277 level in Y221G mutant.

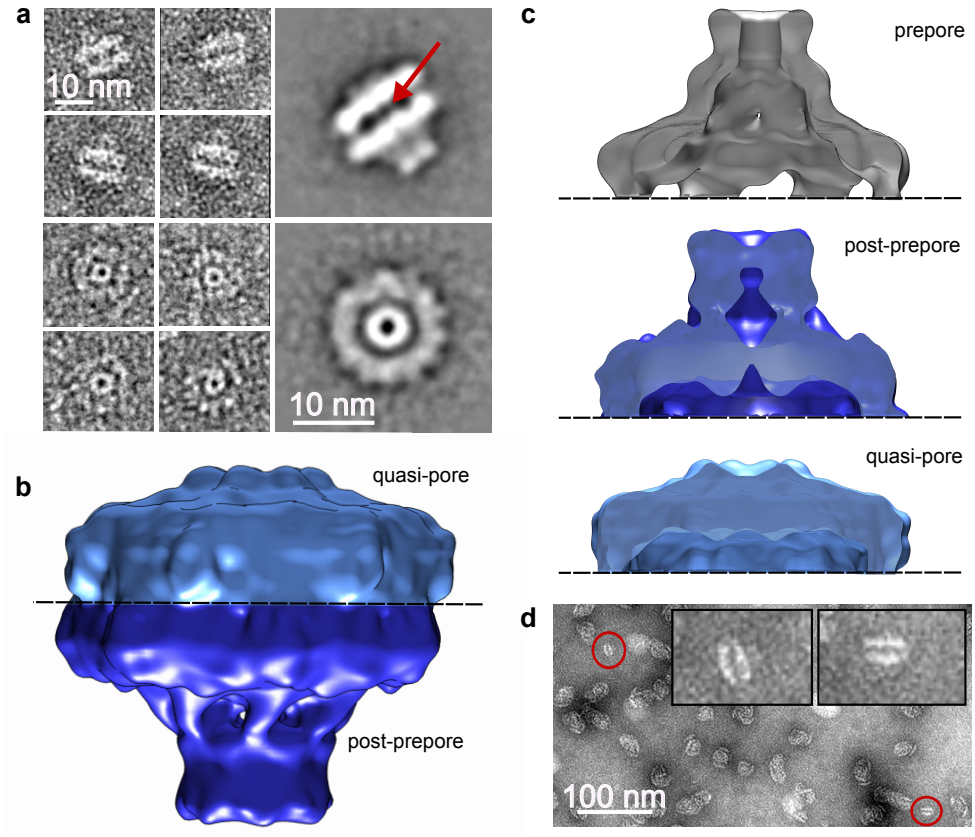


Figure 4. Multiple stages towards pore formation revealed by the K246C–E258C mutant. (a) Side (top) and top (bottom) EM views of negative-stained heptameric K246C–E258C mutant samples and class averages there of. A stem structure, highlighted by a red arrow, is visible between the two asymmetric heptamers. (b) Cryo-EM map (18.3 Å resolution, Supplementary Fig. 2b,c) of the asymmetric K246C–E258C oligomeric particles corresponding to two heptamers with different configurations (i.e. post-prepore in dark blue and quasi-pore in light blue) derived from cryo-EM particles (see Supplementary Fig. 2b,c). (c) Comparison of cryo-EM maps of Y221G heptamer (prepore in gray, Fig. 2b), and the post-prepore and quasi-pore conformations extracted from K246C–E258C heptamer. A wide cavity is observed in the prepore lumen, which gradually fills up as the protein proceeds to the post-prepore and then quasi-pore. Transition of the post-prepore

to the quasi-pore involves a collapse of the structure. (d) Negative stain of a WT heptameric aerolysin sample after centrifugation. Non-aggregated asymmetric double heptamers are circled in red and magnified in the insets.

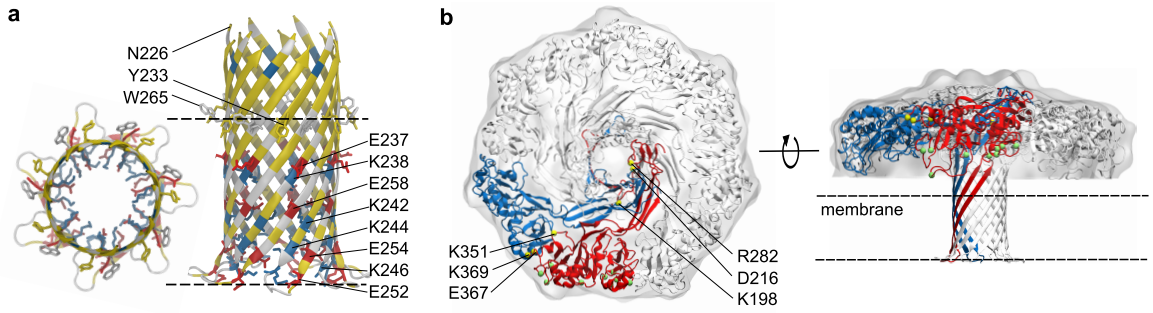


Figure 5. Near-atomistic model structure of the wild-type aerolysin pore. (a) Model of the aerolysin ß-barrel, taking into account the boundaries established by cysteine-scanning analysis and planar lipid bilayer studies (Supplementary Fig. 8a). An aromatic belt (Y233, W265) is ideally positioned to anchor the structure at the lipid head group-acyl chain boundary. The positions of charged residues pointing into the barrel lumen are also indicated (see Supplementary Fig. 8b). (b) Side and top views of the model of the membrane-inserted pore of aerolysin showing the arrangement of the seven monomers in the disk-like structure of the cryo-EM map of the K246C–E258C quasi-pore. Positions are shown for the residues at the monomer-monomer interface that we experimentally showed to be involved in oligomerization (Fig. 2b,d and Supplementary Fig. 4). The position of the lipid bilayer, consistent with conformation of wt aerolysin in nanodiscs (Supplementary Fig. 7a), preliminary cryo-electron tomography (data not show) and MD simulation of the barrel (Supplementary Fig. 8b), is represented by the dashed lines.

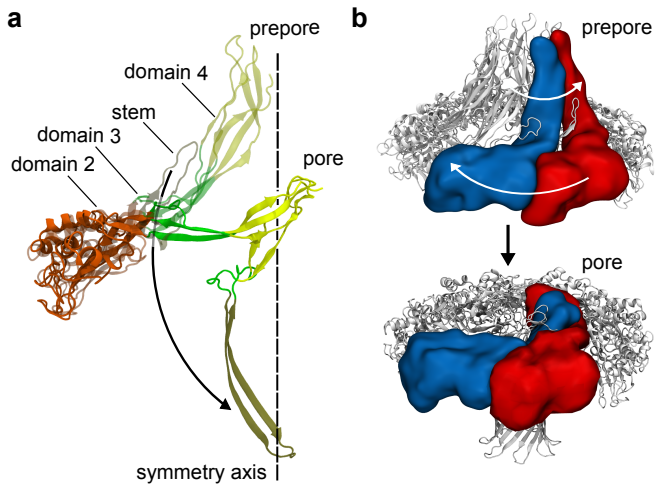


Figure 6. Swirling mechanism promoting transition between the prepore and the pore conformation. (a) Ribbon representations of aerolysin monomer as observed in the model of the prepore (transparent) and the membrane-inserted (solid) state. For sake of clarity domain 1 is not shown. During the prepore to pore transition state, domains 3 (green) and 4 (yellow) rotate and flatten to a position almost parallel to the membrane plane. During this transition, the pre-stem loop is extracted and flips around to form the transmembrane β-barrel (see Supplementary Video 1). (b) Ribbon and space-filled (for two representative monomers) representation of the prepore and the membrane-inserted state illustrating the overall swirling movement undergone by the complex, the change in the relative position of domains 4 in the heptamer, and the increase in monomer-monomer contact surface (see Supplementary Video 1). The blue space-filled monomer has the same orientation as in panel (a).

ONLINE METHODS

Toxin production and activity assays. His-tagged protein production and purification was performed as described[13]. For crystallization, pro-aerolysin Y221G was however produced in the absence of any tag in *Aeromonas salmonicida* CB3 strain, a protease deficient mutant (Rif^R, Kn^R) and purified as described[33] [8]. The Y221G mutant was further purified by size exclusion chromatography using a Superdex^{TM} 200 10/300 GL column (GE Healthcare) equilibrated in Tris-HCL 20mM pH 7.4, NaCl 150mM. The recovered Y221G mutant was concentrated by centrifugation in 30 kDa molecular-weight cutoff Vivaspin centrifugal concentrator (VivaScience, Sartorius Group) to a concentration of 5 mg/ml at 4°C. Mutagenesis was carried out using QuickChange II Site Directed Mutagenesis Kit from Agilent Technologies. The activity of all mutants was tested by hemolysis[13]. *In vitro* oligomerization was performed as described[5]. Briefly, the protein was activated in 20mM Tris 150mM NaCl pH 8.2 with insoluble trypsin (Sigma *T1763*) for 2-4h at 4 °C. The trypsin was removed by centrifugation and the activated toxin was dialyzed against 20mM Hepes 50mM NaCl pH 7. Oligomerization was performed at room temperature with (WT reconstitution) or without (single particles analysis) 0.5% C8E5 detergent[5].

X-ray crystallography. Crystallization trials were performed by using the vapour-diffusion method and the hanging-drop technique. The crystals typically appeared under the following conditions: 2μl of solution protein at the concentration 2mg/ml was mixed with 2μl of the reservoir solution containing 9%-11% PEG4000, sodium acetate 100mM pH 5.4-5.8. The drops were let to equilibrate and kept at 18°C. After one week the crystals grew as large thin plates with dimensions of 0.25 x 0.10 x 0.05 mm^3. The crystal belong to the orthorhombic space group $P2_12_12_1$ with unit cell dimensions of a=72.62 Å, b=92.43Å, c=169.05Å. The asymmetric unit accommodates two monomers, corresponding to a crystal

packing parameter (V_M) of 2.7 A^3/Da and a solvent content of 54%. A first X-ray diffraction data set was recorded at room temperature (291 K) on a single crystal at the Swiss and Norvegian Beamline (BM1) at the European Synchrotron Radiation Facility (Grenoble, France). The beam-line was tuned at the wavelength of 0.82 Å and a Marresearch Image Plate detector was used. Raw diffraction images were indexed and integrated with *MOSFLM*[34] and scaled with SCALA [35] within the CCP4 program suite (1994). The structure of the Y221G proaerolysin was determined at 2.86 Å by molecular replacement using *MOLREP* [36] and the wild-type proaerolysin structure solved at 2.88 Å as a search model (1PRE). Prior to the molecular replacement procedure, the tyrosine residue at the position 221 was kept (Supplementary Fig. 1a).

Rounds of refinement performed with CNS[37] were interchanged with map inspection and manual rebuilding of the model with the molecular graphics program O[38]. In the first cycles of refinement, the domain 4 (residues 196-223, 275-298, 410-470) was introduced and reconstructed in the electron density map. The presence of a glycine residue at position 221 on both chain A and B was also confirmed on both subunits of the Y221G proarolysin dimer. The program MolProbity [39] was used to analyse the stereochemical quality of the final refined model and it revealed that 90.2% of the residues are located in the most-favored regions of the Ramachandran plot. Supplementary Table 1 summarizes all relevant statistics of data collection and refinement. A second X-ray diffraction data set at a higher resolution of 2.2 Å was recorded on a single crystal cryocooled at 100 K at the beamline x06sa at the Swiss Light Source located in the Paul Scherrer Institute (Villigen, Switzerland). The beam-line was tuned at the wavelength of 0.95 Å and equipped with a Marresearch CCD detector. The crystals were briefly soaked in a cryoprotectant solution having the same composition as the mother liquor but with a slight increase of PEG 4000 to 14% and containing PEG 400 as cryoprotectant. After this immersion in the cryoprotectant solution, the crystal was flash-cooled by plunging into liquid nitrogen. Refinement of the model against the data at 2.2 Å was done with CNS

and finished with Refmac5[40] (Supplementary Table 1). The stereochemical quality of the final refined model was assessed with MolProbity and it revealed that 95.5% of the residues are located in the most-favored regions of the Ramachandran plot. Finally, the Y221G crystals were soaked in a solution containing mannose-6-phosphate (M6P) at a concentration of 53 mM. The crystals were cryoprotected using 20% PEG400 and cryocooled at 100K. A third and a final data set at a resolution of 2.5 Å was collected on a single crystal at the beam-line x06sa at the SLS. The wavelength was set at 0.979 Å and a pixel detector (PILATUS 6M, DECTRIS Ltd., Switzerland) was used. Likewise, for the refinement procedure of the structure Y221G proaerolysin complexed with mannose-6-phosphate, the programs Refmac5 and O were used (Supplementary Fig. 3a and Table 1). The analysis of stereochemical quality of the final refined model done with MolProbity showed that 92.9% of the residues are located in the most-favored regions of the Ramachandran plot.

Electron microscopy. A 4 μl drop of aerolysin sample at ~1mg/ml was applied on a formvar-carbon grid (Electron Microscopy Sciences FCF-200-Cu) for 30s and then blotted. The grid was subsequently washed with three drops of ultrapure water and stained with 2% Uranyl acetate. Grids were imaged on a Tecnai Spirit at 80kV. Samples for cryo-EM were prepared on holey carbon grids (Quantifoil Micro Tools GmbH Jena, Germany). Grids were plunge-frozen in liquid nitrogen cooled liquid ethane with a FEI Vitrobot and imaged on a FEI Titan Krios at 300kV equipped with a US4000 CCD.

Image processing. Image processing and 3D refinement was performed using EMAN2 [41] according to the eman2 workflow. Particles were manually selected and CTF corrected followed by phase-flipping (11000 particles for the K246C-E258C mutant and ~5000 for the Y221G). A reduced particle set (~1000 particles) was used to generate the initial 2D reference free class averages used to obtain an initial 3D model applying a c7 (K246C–E258C mutant) or d7 (Y221G

mutant) symmetry. The initial model was refined through 5 iterations using the imposed symmetry and 3'' (Y221G) or 5'' (K246C-E258C) angular increments while using the default automask 3D function implemented in EMAN2. The obtained model was lowpass filtered using a cutoff frequency of 1/20 Å and was used as initial model in obtaining a high resolution model using the full particle set during 5 iterations using an angular step of 1.5''. The resolution of the resulting models was assessed by splitting the data into even and odd halves and computing semi-independent reconstructions using the implemented e2eotest in EMAN2 (Supplementary Fig. 2c)[42]. The final maps used for molecular assembly and dynamics were lowpass-filtered at 20 Å. All EM maps presented in throughout the paper were generated using Chimera[43].

Planar lipid bilayers. Planar lipid bilayers were performed as previously described [13] using a Warner Instrument Corp. planar lipid bilayer setup. Currents were recorded at either ±25 mV in a Faraday cage, amplified (BC-525c, Warner Instrument Corp.), low-pass filtered with an eight-pole Bessel filter at 4 kHz (LPF-8, Warner Instrument Corp.), A-D converted (DigiData 1200, Axon Instruments, Union City, CA) and sampled at 10 kHz by computer (pClamp 8.0, Axon Instruments).

Aerolysin structural models and molecular simulations. The structures of WT and Y221G pro-aerolysin (pdb 1PRE and 3C0N) were used to prepare model systems for molecular dynamics simulations. In both crystal structures the region 423 to 439 is missing. This corresponds to a flexible loop connecting the CTP to the rest of the protein, and containing the site where cleavage takes place during aerolysin activation (420-427). We assumed that these structures would mimic the aerolysin structure after cleavage, i.e. CTP no longer covalently connected to the protein, but still bound to it. Domain 1 is known to act as an independent folding unit[25], and is connected by a long random coil region to the bigger lobe. For this reason, we decided to remove it from every system in order to reduce the

system size and therefore speedup the molecular dynamics computation time. All simulations of monomeric aerolysin (i.e. WT-CTP, WT, Y221G-CTP, Y221G, as well as WT with displaced stem) were solvated in a rectangular box of TIP3P water and neutralized by NaCl. Simulations were run using NAMD [44] MD engine, using Amber99SB force field[45], SHAKE algorithm on all the bonds within hydrogen and heavy atoms, and Particle-mesh Ewald treating the electrostatic interactions in periodic boundary conditions. We chose an integration step of 2 fs. Temperature has been controlled by means of Langevin forces, using a damping constant of 1 ps^{-1}. Principal Component Analysis (PCA) in Fig. 1 and Supplementary Fig. 1 was performed on 2000 frames containing Cα coordinates of WT and Y221G simulations (one frame every 100 ps of simulation) using Gromacs tools[46]. Projection on the same eigenspace of WT-CTP, WT, Y221G-CTP and Y221G trajectories (Fig. 1, and Supplementary Fig. 1 and 4) was obtained by performing a PCA on the Cα coordinates of a pseudo-trajectory containing the last 100 ns of every simulation. In order to steer the stem region of aerolysin WT, a steering force was applied on α-carbon atoms in region E236 to A265. The rotation axis was identified as the normal of the plane identified within residues S236, N262 and G251. S235 was selected as the rotation hinge. A biasing force was applied so that the selected atoms would rotate of 180° in 1 ns. The resulting structure was simulated for 85 ns. All analysis and rendering of protein structures were performed with VMD[47].

In order to better characterize the alignment of the prepore with respect to the lipid bilayer, we modeled one of aerolysin smallest human receptors, the GPI-anchored protein CD52 (Uniprot entry P31358). For a review on GPI-anchors, see ref. [48]. Remarkably, when GPI-anchored, CD52 is just a 12 amino acid-long peptide. No structure of CD52 is available, the peptide is however predicted as disordered (data not shown). For this reason, we modeled CD52 region G25-S36 as a random coil. Importantly, residue N27 is N-glycosylated[49]. We modeled the N-linked sugar, as well as the GPI-anchor glycan core, using Glycam parameters[50]. The inositol of the anchor and its phospho-ethanolamine side

chains (typical for CD52 anchors), were parameterized via quantum mechanical calculations performed using Gaussian 09[51]. The complete receptor (glycosylated CD52 anchored by its C-terminal to the GPI-anchor) was finally assembled, and inserted into a POPC bilayer model. The resulting system was solvated with TIP3P water, and was equilibrated using a MD protocol similar to that reported above. The resulting simulation shows that the GPI-anchor never stretches far from the membrane. In fact, protein and GPI sugars tend to lie on the lipid polar head groups. This is not surprising, since CD52 is rich in serine and threonine residues (T29, S30, T32, S33, S34, S36), which are likely to interact with the membrane. The last snapshot from our GPI-anchor simulation was extracted in order to model aerolysin prepore binding (see Fig. 2c and Supplementary Figure 3).

Molecular assembly. WT and Y221G MD simulations were used to prepare a structural database to build assemblies of the prepore and pore complexes. Predictions of multimeric assemblies were performed using a new kind of particle swarm optimization (PSO) algorithm[20] implemented in a novel framework called *parallel optimization workbench* - *pow*^{er} (available at http://lbm.epfl.ch) and recently applied to other multimeric complexes[52]. This technique allows to quickly find a reasonable prediction for a multimeric structure arrangement on the basis of an ensemble of monomeric protein conformations representative of the protein conformational space, and experimental measures acting as search restraints. For the prepore and pore structure prediction, a PSO search was launched 3 times, for 200 iterations, using 80 particles. Eigenvectors describing at least 90% of the protein movement were automatically selected to describe the proteins native flexibility. To score the quality of an assembly (namely, its fitness), geometric contributions extracted from the cryo-EM maps were taken into account (i.e. Y221G prepore had to be 85±5 Å high, 150±5 Å wide, with a pore radius of 5±1 Å; the WT pore had to be less than 45 Å high, 150±5 Å wide, with a

pore radius of 5±1 Å and residue G270 had to be located less than 7 Å from the pore central axis). A simple energetic term, meant to avoid configurations containing severe clashes, was computed as 9-6 Lennard-Jones potential within the Cα and Cβ of two neighboring monomers.

Solutions having an RMSD smaller than 5 Å were clustered. While for prepore prediction six representative clusters were returned, in quasi-pore prediction no solution matched the filtering criteria (fitness smaller than 0). Still, the model having the best fitness respected all the provided geometric constraints, and presented just few mild backbone clashes. This was therefore accepted as representative model. Prepore and pore selected models were docked into their respective cryo-EM maps using Situs[53] and ranked according to their cross-correlation coefficient (CCC) with respect to the density map. In order to perform this comparison, density maps having the same resolution of the cryo-EM were generated for every model. The CCC within the two maps was subsequently computed. The best scoring prepore and quasi-pore models were finally optimized by means of a Molecular Dynamics Flexible Fitting (MDFF)[21] protocol as implemented in NAMD. In this phase, the protein secondary structure was restrained in order to avoid over-fitting. For both structures, 5 ns simulation were performed *in vacuo* at 300 K and 1 atm, with SHAKE restraining all bonds and a dielectric constant of 80. Unbiased MD was run to finally relax the systems.

For the modeling of the complete pore structure, its β-barrel (region S229 to G269) was designed separately using α-hemolysin barrel's backbone as template. The resulting model was inserted in a POPC bilayer, solvated with TIP3P water, ionized with 0.15 M NaCl and equilibrated with MD at 300 K for 1 ns (Supplementary Fig. 7). This modeled β-barrel was connected to the extracellular-membrane assembly obtained with *pow*^{er}. The obtained system was finally refined by means of a MD protocol. Atoms in regions 268-283 and 214-231 were minimized while the other atoms were kept fixed. Subsequently, a simulated annealing protocol in implicit water (dielectric equal to 80) was performed on region 268-283 and 214-231, and the system was finally equilibrated at 300 K for

1 ns. A 100 frames morphing within the prepore and the quasi-pore model was computed using Chimera[54] (Supplementary Video 1).

The coordinates and the structure factors amplitudes have been deposited in the Protein Data Bank under the accession codes of 3C0M for Y221G proaerolysin at 2.88 Å, 3C0N for Y221G proaerolysin at 2.2 Å and 3C0O for Y221G proaerolysin complexed with M6P at 2.5 Å.

References

1 Bischofberger, M., Gonzalez, M. R. & van der Goot, F. G. Membrane injury by pore-forming proteins. *Curr Opin Cell Biol* **21**, 589-595 (2009).

2 Iacovache, I., Bischofberger, M. & van der Goot, F. G. Structure and assembly of pore-forming proteins. *Curr Opin Struct Biol* **20**, 241-246 (2010).

3 Szczesny, P. *et al.* Extending the aerolysin family: from bacteria to vertebrates. *PLoS One* **6**, e20349 (2011).

4 Parker, M. W. *et al.* Structure of the *Aeromonas* toxin proaerolysin in its water-soluble and membrane-channel states. *Nature* **367**, 292-295 (1994).

5 Iacovache, I. *et al.* Dual chaperone role of the C-terminal propeptide in folding and oligomerization of the pore-forming toxin aerolysin. *PLoS pathogens* **7**, e1002135 (2011).

6 Hong, Y. *et al.* Requirement of N-glycan on GPI-anchored proteins for efficient binding of aerolysin but not Clostridium septicum a-toxin. *EMBO J* **21**, 5047-5056 (2002).

7 Wilmsen, H. U., Leonard, K. R., Tichelaar, W., Buckley, J. T. & Pattus, F. The aerolysin membrane channel Is formed by heptamerization of the monomer. *EMBO Journal* **11**, 2457-2463 (1992).

8 Tsitrin, Y. *et al.* Conversion of a transmembrane to a water-soluble protein complex by a single point mutation. *Nat Struct Biol* **9**, 729-733 (2002).

9 Abrami, L., Fivaz, M., Glauser, P.-E., Parton, R. G. & van der Goot, F. G. A pore-forming toxin interact with a GPI-anchored protein and causes vacuolation of the endoplasmic reticulum. *J. Cell Biol.* **140**, 525-540 (1998).

10 Nelson, K. L., Raja, S. M. & Buckley, J. T. The GPI-anchored surface glycoprotein Thy-1 is a receptor for the channel-forming toxin aerolysin. *J. Biol. chem.* **272**, 12170-12174 (1997).

11 MacKenzie, C. R., Hirama, T. & Buckley, J. T. Analysis of receptor binding by the channel-forming toxin aerolysin using surface plasmon resonance. *J Biol Chem* **274**, 22604-22609 (1999).

12 Fivaz, M. *et al.* Differential sorting and fate of endocytosed GPI-anchored proteins. *EMBO J* **21**, 3989-4000 (2002).

13 Iacovache, I. *et al.* A rivet model for channel formation by aerolysin-like pore-forming toxins. *The EMBO Journal* **25**, 457-466 (2006).

14 Melton, J. A., Parker, M. W., Rossjohn, J., Buckley, J. T. & Tweten, R. K. The identification and structure of the membrane-spanning domain of the Clostridium septicum alpha toxin. *J Biol Chem* **279**, 14315-14322 (2004).

15 Lasker, K. *et al.* Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 1380-1387 (2012).

16 Alber, F. *et al.* The molecular architecture of the nuclear pore complex. *Nature* **450**, 695-701 (2007).

17 Seidelt, B. *et al.* Structural Insight into Nascent Polypeptide Chain-Mediated Translational Stalling. *Science* **326**, 1412-1415 (2009).
18 Mueller, M., Grauschopf, U., Maier, T., Glockshuber, R. & Ban, N. The structure of a cytolytic alpha-helical toxin pore reveals its assembly mechanism. *Nature* **459**, 726-730 (2009).
19 Shatursky, O. *et al.* The mechanism of membrane insertion for a cholesterol-dependent cytolysin: a novel paradigm for pore-forming toxins. *Cell* **99**, 293-299 (1999).
20 Degiacomi, M. T. & Dal Peraro, M. Macromolecular symmetric assembly prediction using swarm intelligence dynamic modeling. *Structure*, doi:10.1016/j.str.2013.05.014 (2013).
21 Trabuco, L. G., Villa, E., Mitra, K., Frank, J. & Schulten, K. Flexible Fitting of Atomic Structures into Electron Microscopy Maps Using Molecular Dynamics. *Structure* **16**, 673-683 (2008).
22 Zhao, G. *et al.* Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature* **497**, 643-646 (2013).
23 Abrami, L. *et al.* The glycan core of GPI-anchored proteins modulates aerolysin binidng but is not sufficient: the polypeptide moiety is required for the toxin-receptor interaction. *FEBS Lett* **512**, 249-254 (2002).
24 Treumann, A., Lifely, M. R., Schneider, P. & Ferguson, M. A. J. Primary structure of CD52. *J. Biol. Chem.* **270**, 6088-6099 (1995).
25 Lesieur, C. *et al.* Increased stability upon heptamerization of the pore-forming toxin aerolysin. *J Biol Chem* **274**, 36722-36728 (1999).
26 Ritchie, T. K. *et al.* Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods in enzymology* **464**, 211-231 (2009).
27 Song, L. *et al.* Structure of Staphylococcal a-hemolysin, a heptameric transmembrane pore. *Science* **274**, 1859-1866 (1996).
28 Melton-Witt, J. A., Bentsen, L. M. & Tweten, R. K. Identification of functional domains of Clostridium septicum alpha toxin. *Biochemistry* **45**, 14347-14354 (2006).
29 De, S. & Olson, R. Crystal structure of the Vibrio cholerae cytolysin heptamer reveals common features among disparate pore-forming toxins. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7385-7390 (2011).
30 Akiba, T. *et al.* Crystal structure of the parasporin-2 Bacillus thuringiensis toxin that recognizes cancer cells. *J Mol Biol* **386**, 121-133 (2009).
31 De Colibus, L. *et al.* Structures of lysenin reveal a shared evolutionary origin for pore-forming proteins and its mode of sphingomyelin recognition. *Structure* **20**, 1498-1507 (2012).
32 Bischofberger, M., Iacovache, I. & van der Goot, F. G. Pathogenic pore-forming proteins: function and host response. *Cell host & microbe* **12**, 266-275 (2012).

33 Buckley, J. T. Purification of cloned proaerolysin released by a low protease mutant of *Aeromonas salmonicida*. *Bioch. Cell Biol.* **68**, 221-224 (1990).
34 Leslie, A. G. The integration of macromolecular diffraction data. *Acta crystallographica. Section D, Biological crystallography* **62**, 48-57 (2006).
35 Evans, P. Scaling and assessment of data quality. *Acta crystallographica. Section D, Biological crystallography* **62**, 72-82 (2006).
36 Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *Journal of Applied Crystallography* **30**, 1022-1025 (1997).
37 Brünger, A. T. The Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-474 (1992).
38 Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **A47**, 110-119 (1991).
39 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **66**, 12-21 (2010).
40 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica. Section D, Biological crystallography* **53**, 240-255 (1997).
41 Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *Journal of structural biology* **157**, 38-46 (2007).
42 van Heel, M. *et al.* Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys* **33**, 307-369 (2000).
43 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25** (2004).
44 Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-1802 (2005).
45 Case, D. *et al.* AMBER 11. *University of California, San Francisco* (2010).
46 Neres, J. *et al.* Structural basis for benzothiazinone-mediated killing of Mycobacterium tuberculosis. *Science translational medicine* **4**, 150-121 (2012).
47 Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38 (1996).
48 Paulick, M. G. & Bertozzi, C. R. The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. *Biochemistry* **47** (2008).
49 Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-147 (2002).
50 Peters, J. H. & de Groot, B. L. Ubiquitin dynamics in complexes reveal molecular recognition mechanisms beyond induced fit and conformational selection. *PLoS computational biology* **8**, e1002704 (2012).

51 Plaxco, K. W. & Gross, M. Protein complexes: the evolution of symmetry. *Current biology : CB* **19**, R25-26 (2009).
52 Hofmeyer, T. *et al.* Arranged sevenfold: structural insights into the C-terminal oligomerization domain of human C4b-binding protein. *J Mol Biol* **425**, 1302-1317 (2013).
53 Wriggers, W. Using Situs for the integration of multi-resolution structures. *Biophysical reviews* **2**, 21-27 (2010).
54 Manina, G. *et al.* Biological and structural characterization of the Mycobacterium smegmatis nitroreductase NfnB, and its role in benzothiazinone resistance. *Mol Microbiol* **77**, 1172-1185 (2010).

SUPPLEMENTARY INFORMATION

Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism

Matteo T. Degiacomi[1#1], Ioan Iacovache[2#2], Lucile Pernot[2],

Mohamed Chami[3], Misha Kudryashev[3], Henning Stahlberg[3],

F. Gisou van der Goot[2*] and Matteo Dal Peraro[1*]

[1] Institute of Bioengineering and [2] Global Health Institute
School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne - EPFL
CH-1015 Lausanne, Switzerland, [3] Center for Cellular Imaging and NanoAnalytics (C-CINA), Biozentrum, University Basel, Mattenstrasse 26, CH-4058 Basel, Switzerland

[#] These authors contributed equally to this work

* Correspondence to be addressed to G.F.V.D.G. or M.D.P. (co-senior authors)

Institute of Bioengineering / Institute of Global Health
Ecole Polytechnique Fédérale de Lausanne
Station 19, 1015 Lausanne, Switzerland
E-mails: matteo.dalperaro@epfl.ch, gisou.vandergoot@epfl.ch

[1] Current address: Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford, OX1 3QZ
[2] Current address: Institute of Anatomy, University of Bern, Baltzerstrasse 2, 3012 Bern

SUPPLEMENTARY RESULTS

SUPPLEMENTARY TABLES

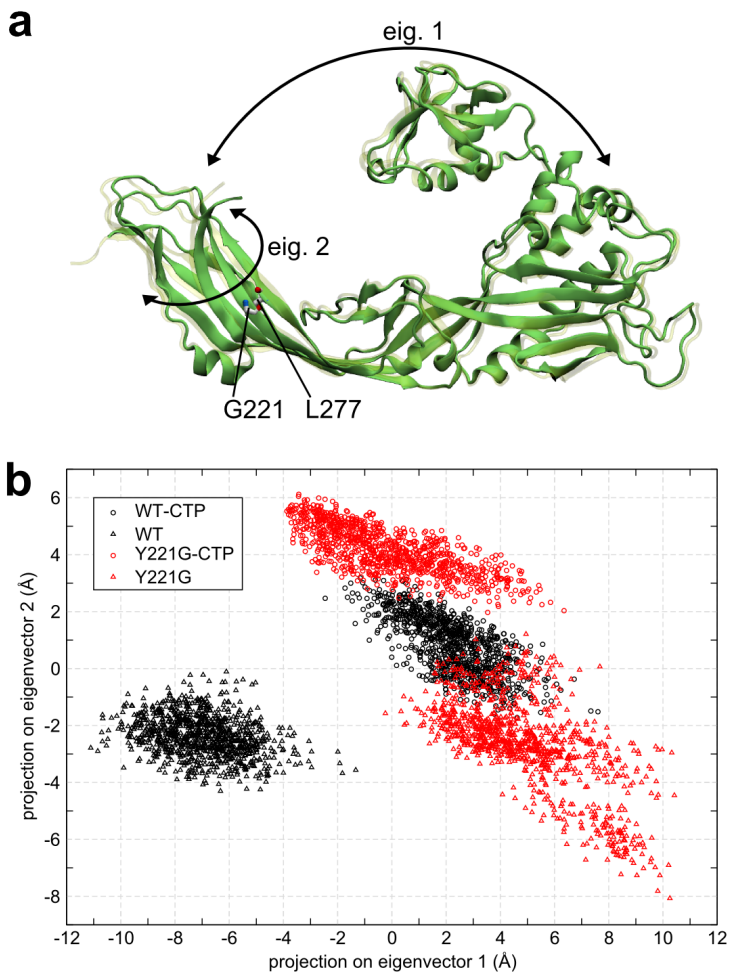
Supplementary Table 1. X-ray data collection and refinement statistics. The number in parentheses represent the highest resolution shell: Y221G (291K) 3.0 Å - 2.86 Å, Y221G (100K) 2.32 Å - 2.2 Å, Y221G-M6P 2.64 Å- 2.5 Å.

	Y221G (291K)	Y221G (100K)	Y221G-M6P (100K)
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	72.62, 92.43, 169.05	70.36, 89.77, 166.62	70.95, 90.78, 165.80
A, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å) (1)	44.6 - 2.88 (3.0 - 2.88)	64.8 - 2.2 (2.32 - 2.2)	61.2 - 2.5 (2.64 – 2.5)
R_{sym} (%)	8.7 (36.3) (2)	11.3 (49.8)	7.6 (31)
I / σI	10.4 (2.2)	27.0 (8.0)	14.01 (4.1)
Completeness (%)	91.7 (67.4)	100 (100)	99.3 (98.7)
Redundancy	2.8 (2.4)	21.3 (18.2)	4.8 (4.5)
Refinement			
Resolution (Å)	44.6 - 2.88	64.8 - 2.2	61.2 – 2.5
No. reflections	23054	51580	35664
R / R_{free} (%)	20 / 25.2	18.9 / 23.8	20.3 / 27.4
No. of			

protein residues (A/B)	449/449	450/450	450/450
Ligand/ion	-	-	2 / 1
Water	24	417	173
B-factors (Å^2)			
Protein	36.1	31.8	41.7
Ligand/ion	-	-	56 / 54.9
Water	24.0	32.1	33.7
R.m.s. deviations			
Bond lengths (Å)	0.08	0.02	0.022
Bond angles (°)	1.5	1.7	2.0

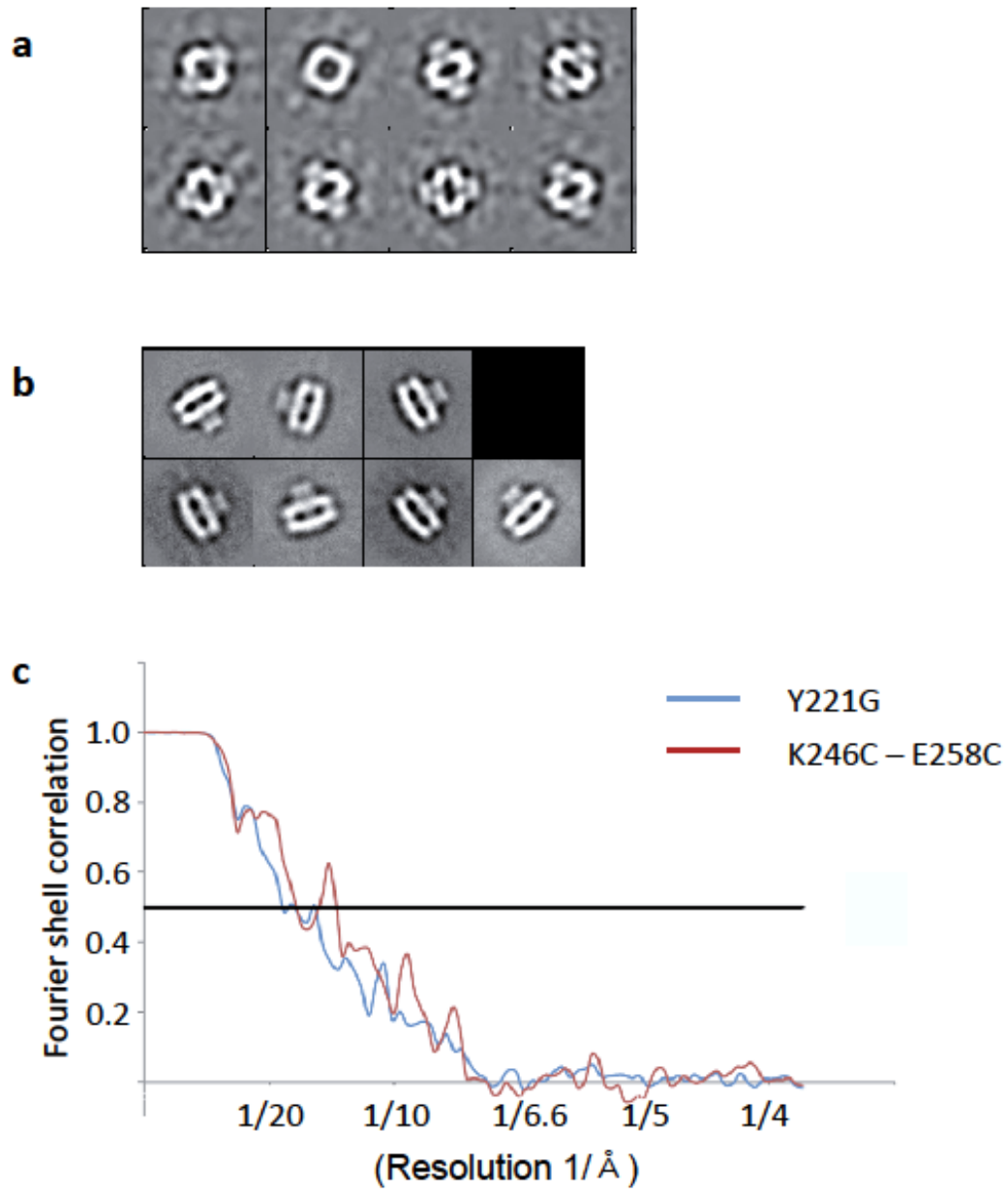
Each reported data set was collected on a single crystal. (1) The numbers in parentheses delimit the highest resolution shell. (2) The value obtained in the highest resolution shell is in parentheses.

SUPPLEMENTARY FIGURES



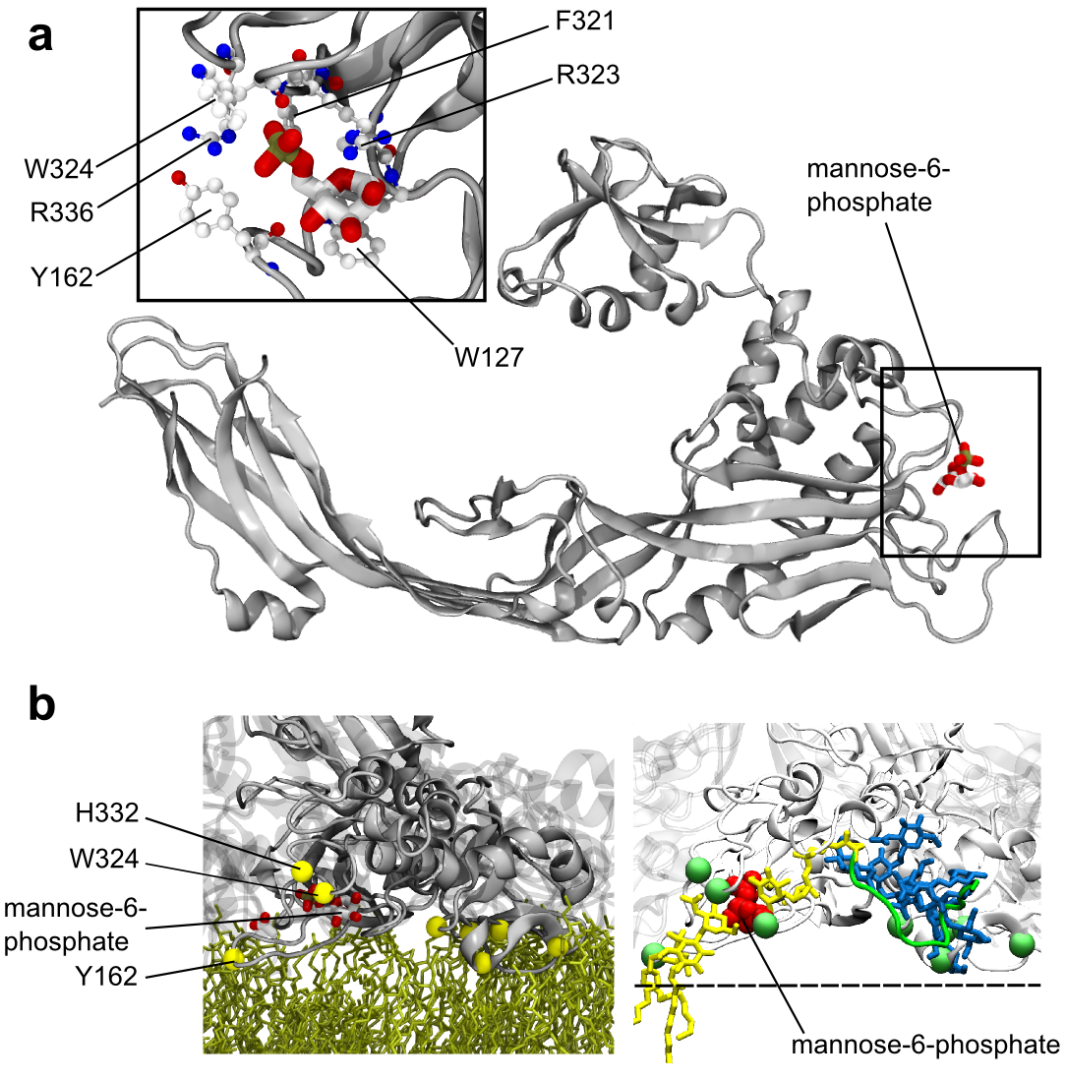
Supplementary Figure 1. Structural flexibility of aerolysin

(a) Crystal structure of pro-aerolysin Y221G (green) superposed to pro-aerolysin WT (transparent). The two structures are nearly identical (backbone RMSD 1.3 Å) with the only exception of the position of residues L277 and G221, which are highlighted in licorice representation (see Table S1 and Figure 3a). Bidirectional arrows describe the proteins' movements along their first and second eigenvector highlighted by Principal Component Analysis (PCA) of MD simulations, see panel (**b**) and Fig. 1b. **(b)** Projection of the last 100 ns of WT-CTP, WT, Y221G-CTP and Y221G trajectories on the first two eigenvectors of their common eigenspace. Eigenvector 1, responsible of 51.4% of total protein motion, represents aerolysin's "opening and closing" movement, while eigenvector 2 (19.9%), represents the twist of domain 4 around its principal axis. Remarkably, upon CTP removal, aerolysin explores a different region of the conformational space determined by the first eigenvector. See Fig. 1b of the main text for a visual representation.



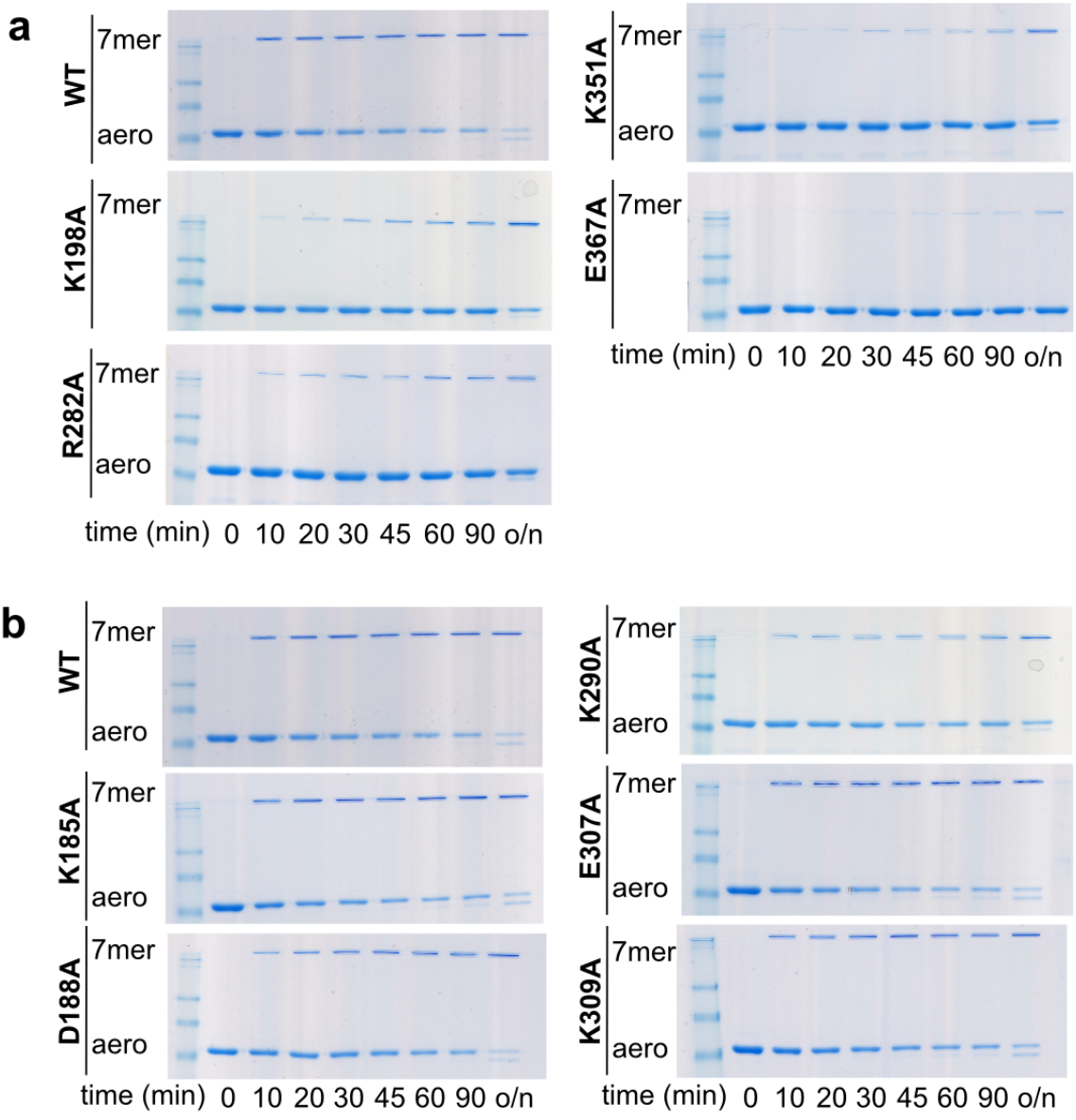
Supplementary Figure 2. Cryo-EM of Y221G and K246C-E258C heptamers

Class averages of Y221G **(a)** and K246C-E258C **(b)** aerolysin heptamers obtained from cryo-EM using the Wiener filtered particles (see Figs. 2a, 4a). **(c)** FSC plots for Y221G (blue) and K246C-E258C (red) as obtained by comparing the even and odd maps generated by EMAN2 *eotest* function. Intersection of the threshold line (black) with each plot shows the estimated resolution, respectively 16.6 Å for Y221G and 18.3 Å for K246C-E258C mutant.



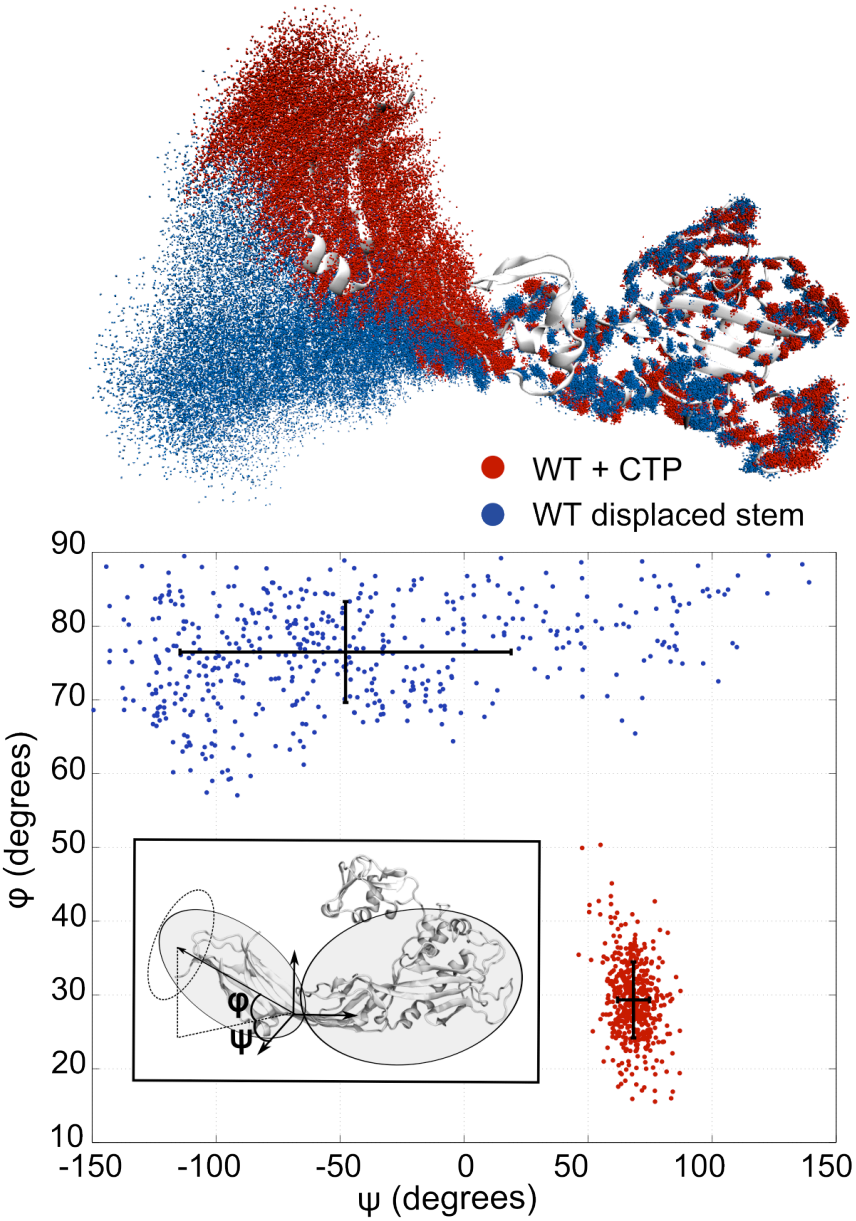
Supplementary Figure 3. Aerolysin prepore bound to GPI receptors

(a) Crystal structure of Y221G pro-aerolysin with mannose-6-phosphate bound to domain 2 binding site. In the inset, a zoom on mannose-6-phosphate, with neighboring residues highlighted in balls and sticks representation (see Fig. 2 and Supplementary Table 1). **(b)** On the left, superimposition of domain 2 of the prepore model with domain 2 of aerolysin X-ray structure having a mannose-6-phosphate molecule trapped in the binding site. The molecule is located in the binding site, close to the membrane. On the right, model of aerolysin prepore bound to the GPI-anchored CD52 (see Fig. 2c). Residues responsible of binding are highlighted in lime color using a space filling representation. Domain 2 binds to the GPI-anchor glycan core (in yellow). The position of mannose-6-phosphate is highlighted in red. Domain 1 binds to N-linked sugars (in blue) bound to the 12 residue-long CD52 (green).

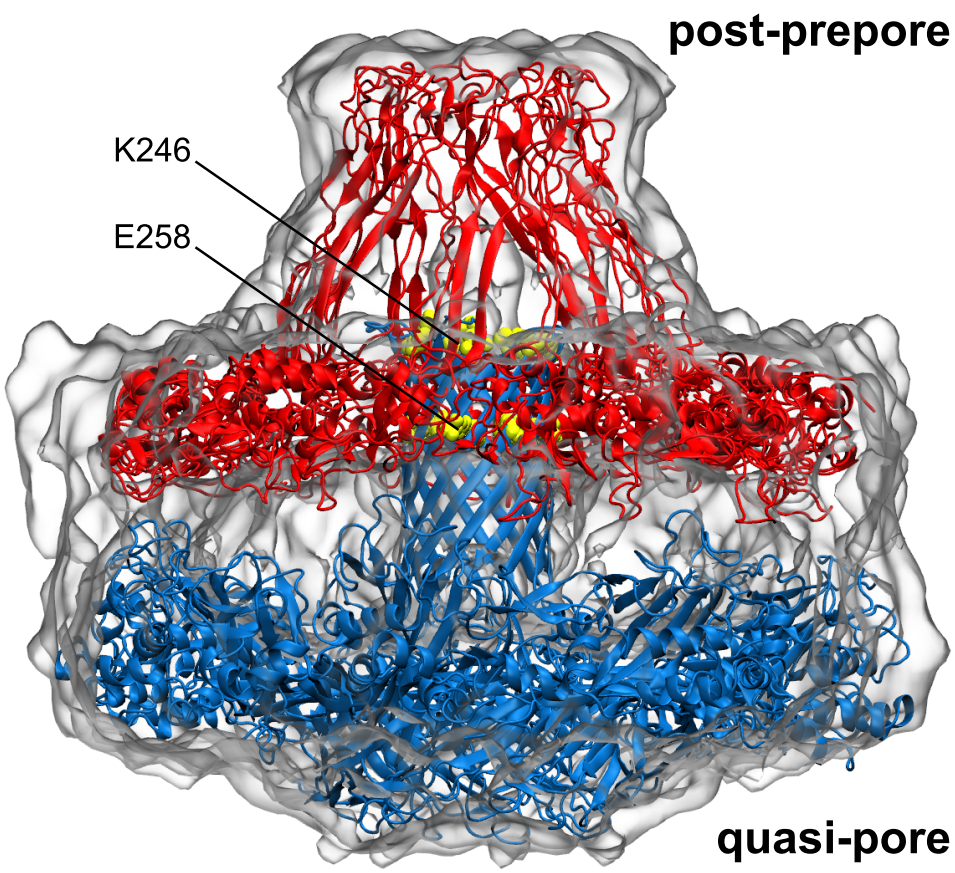


Supplementary Figure 4. Mutagenesis and heptamerization kinetics of WT aerolysin.

(a) Residues predicted to be important for the monomer – monomer interaction in both the pre-pore and the pore models show delayed heptamerization kinetics when mutated to alanine (see Figure 2d and 6b). Aerolysin (WT and mutants) was activated at 4 °C to prevent oligomerization with insoluble trypsin. After removal of trypsin the sample was shifted at room temperature and heptamerization was monitored by SDS-PAGE. **(b)** Residues predicted to be located at the monomer – monomer interface in the pre-pore model but not in the membrane-inserted pore model have similar WT heptamerization kinetics when mutated to alanine (see Figure 2d). Sample preparation as in **(a).**

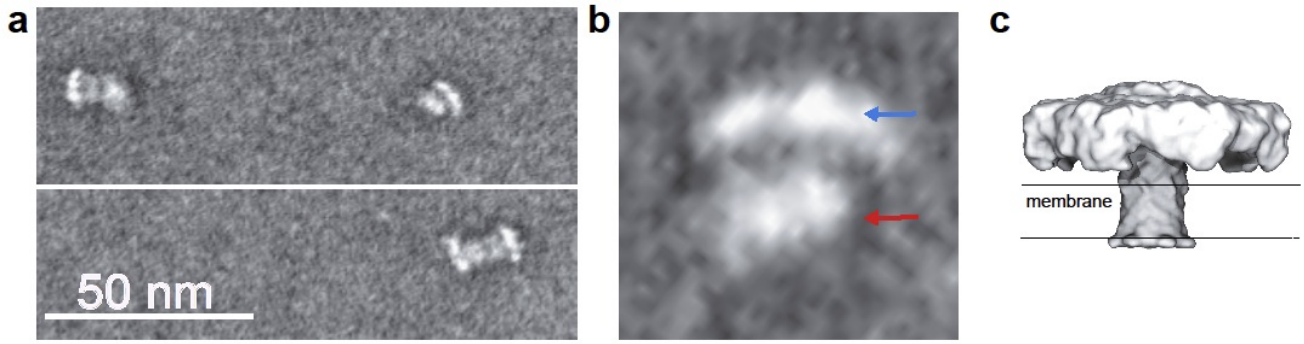


Supplementary Figure 5. Increased flexibility of the wild-type aerolysin upon pre-stem loop extraction. MD simulations reveal that domain 4 flattens down when the pre-stem loop is extracted. Top panel – position of Cα atoms along MD simulations of pro-aerolysin (in red) and aerolysin with displaced stem (in blue) are superimposed on domains 2 and 3. Bottom panel – the abovementioned rearrangement is quantified by plotting the relevant angles ψ and φ (see inset) describing the relative movement of domain 4 with respect to the rest of aerolysin. Notice also the much larger fluctuations of domain 4 observed when the stem is extracted.



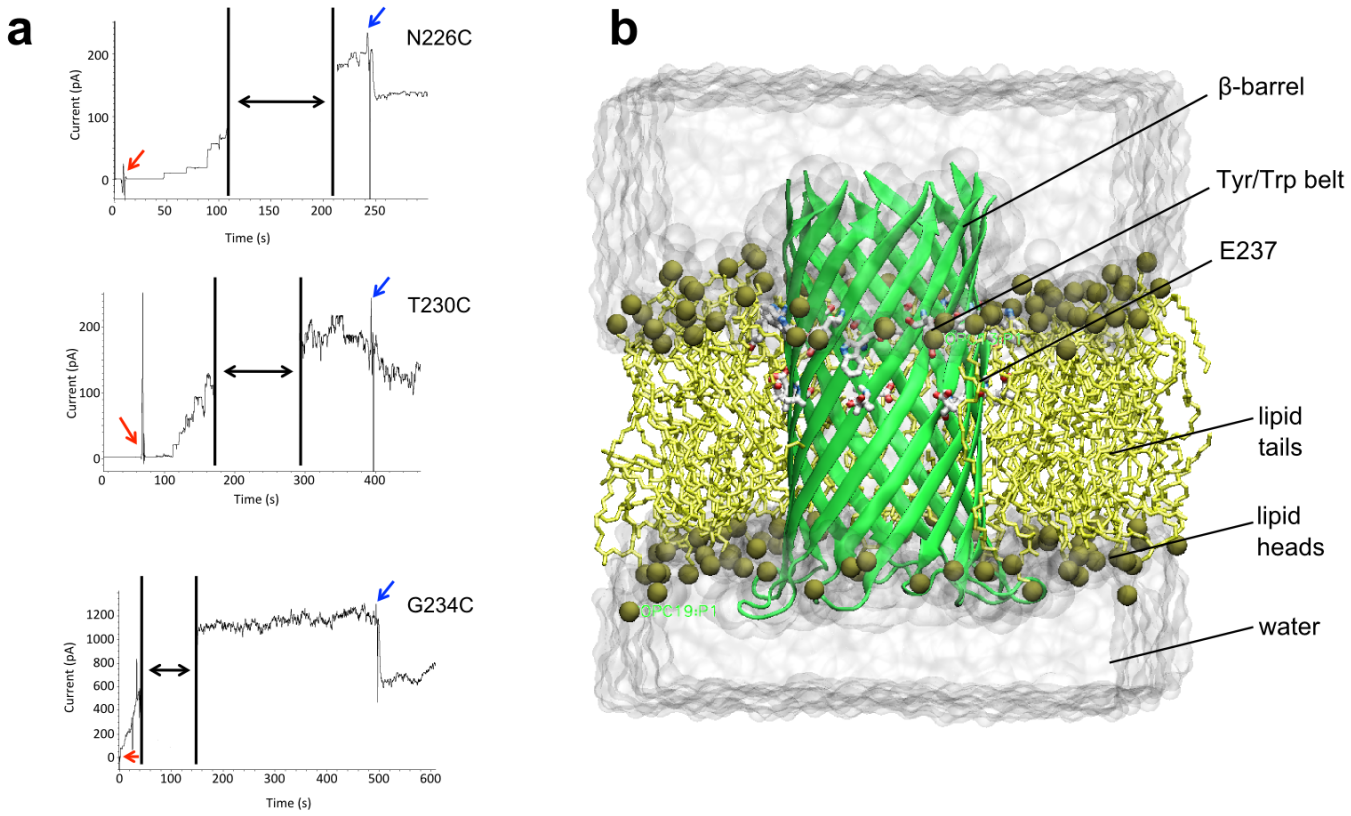
Supplementary Figure 6. Post-prepore and quasi-pore docked in K246C–E258C dimeric EM map.

Post-prepore (in red, without pre-stem, Fig. 2) and quasi-pore (in blue, Fig. 5) can be both docked in K246C-E258C density map (Fig. 4b). Quasi-pore barrel inserts into post-prepore inner cavity. Residues K246 and E258 (in yellow) are neighbors in the pre-stem. Mutation K246C-E258C creates a disulfide bridge impairing pre-stem rearrangement and complete barrel formation. As a result post-prepore cavity can accommodate both post-prepore pre-stem in random coil conformation and quasi-pore's only partially formed ß-barrel.



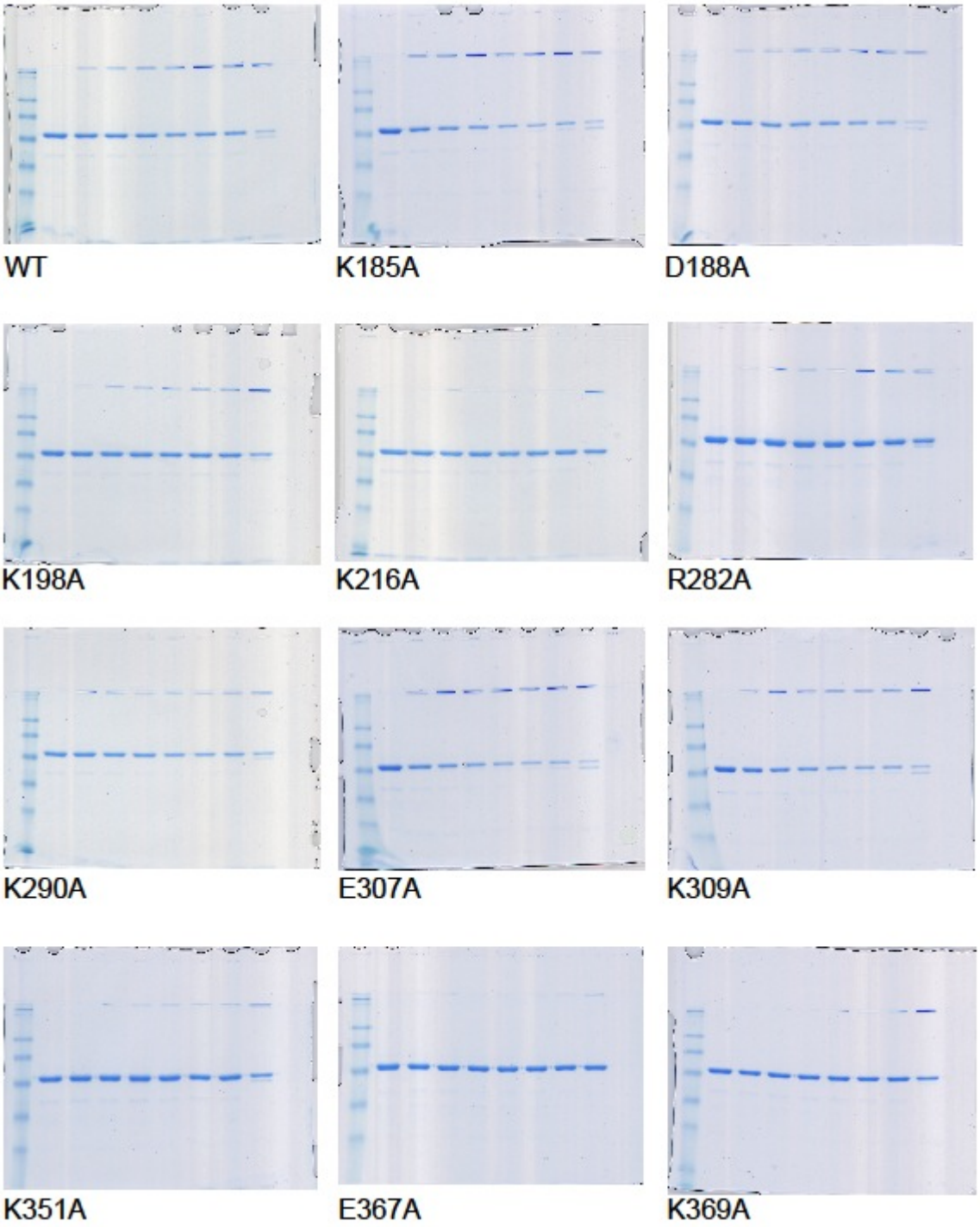
Supplementary Figure 7. Wild-type aerolysin pores in nanodisc membrane environment

(a) Reconstitution of WT aerolysin in POPC:POPE MSP1D1 nanodiscs showing side views of the oligomer. A mixture of aerolysin heptamers, lipids and detergent (C8E5) was incubated for 1h at room temperature prior to detergent removal by biobeads. Sample was briefly centrifuged to remove aggregates and imaged by negative staining. (**b**, **c**) A single aerolysin pore (indicated by a blue arrow) trapped in a nanodisc (red arrow) (**b**) is compared with a surface representation (**c**) of the near-atomistic model of the WT pore state (Fig. 5) at the same scale.



Supplementary Figure 8. Assessing β-barrel length and conformation.

(a) Channels formed by cysteine single point mutants of aerolysin were reconstituted in planar lipid bilayers and formation of membrane-inserted channels was monitored by stepwise increase in current. Red arrow shows the addition of pre-activated aerolysin protein to the *cis* chamber of the planar lipid bilayer setup. After channel formation the buffer in the *cis* chamber was replaced by perfusion (horizontal black arrow) to avoid formation of new channels. Addition of MTS-biotin (blue arrow) results in a drop in current due to partial obstruction of the channel by biotin showing that the residues are facing inside the β-barrel. **(b)** The aerolysin modeled barrel (Fig. 5a) has been inserted in a patch of POPC lipids, solvated with TIP3P water, ionized with 0.15 M NaCl ions, and equilibrated at 300 K and 1 atm. A belt of aromatic residues (Y233 and W265) is located next to lipids polar heads. E237 is the only polar residue exposed to a hydrophobic environment, but could however snorkel (bending the membrane) or, possibly, interact with cholesterols.



Supplementary Figure 9. Raw gel images. Uncut images of all mutagenesis and hetamerization kinetics tests (see Supplementary Fig. 4).

Supplementary Video 1. Aerolysin pore formation via a swirling mechanism

The movie shows a top and a side view of a morphing from the prepore (Fig. 2) to the pore models (Fig. 5). Conversion from one protein arrangement to the other is possible via a swirling mechanism, which can take place without any relevant topological bottlenecks (Fig. 6). Aerolysin domain 1 is not shown in the movie for sake of clarity.